View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by DSpace@MIT

Massachusetts Institute of Technology
Course XX

Thesis Proposal
Doctor of Philosophy

Title:

Design and Evolution of Engineered Biological Systems

Date of Submission:
July 20, 2005

Submitted By: Jason Kelly
31 Ames St. Room 68-558D
Cambridge, MA 02139

Signed: _____

Supervisor: Drew Endy
Assistant Professor
Biological Engineering Department

Signed: _____

Academic Office: Dalia Gabour
Assistant Director, Educational Services
Biological Engineering Department

Signed: _____

Acknowledgements

I would like to thank Tom Knight, Drew Endy, and Randy Rettberg for developing the BioBricks infrastructure on which much of this project is based. Conversations with all three have contributed significantly to the final form of this proposal. In particular, Drew helped me to communicate my ideas in a way that *may* make sense to someone other than myself. Johnathan Goler provided plasmids and strains from the laboratory of Jay Keasling. Frederick Balagadde taught me how to design microfluidic chips, personally fabricated chips for me, and has consistently gone out of his way to help me develop the Sortostat; his help has been invaluable. Steven Quake allowed me to spend time in his laboratory working with Frederick and also founded the Caltech Microfluidic Foundry where my chips are fabricated. All the members of the Endy lab deserve thanks for giving me great feedback regarding this project and for putting up with me in general. Josh Michener is responsible for most of the work described in the preliminary results, and has been an absolutely outstanding UROP student. Funding for this work has been provided in part by the National Science Foundation. Finally, I would like to thank the members of my thesis committee:

Edward F. DeLong
Biological Engineering Department & Department of Civil and Environmental Engineering
MIT

Drew Endy
Biological Engineering Department
MIT

Thomas Knight Jr.
Computer Science and Artificial Intelligence Labroatory
MIT

1 Abstract

To date, engineered biological systems have been constructed via a variety of *ad hoc* approaches. The resulting systems should be thought of as pieces of art. Here, I propose to explore how existing forward engineering approaches might be combined with evolution to make routine the construction of engineered biological systems. I will specify a procedure for construction of biological systems via screening of subcomponent libraries and rational re-assembly. I will develop tools to enable this approach including a high-throughput screening system to measure the input/output function of an arbitrary genetic device. I will apply this approach to construct a collection of ring oscillators and bi-stable switches. Furthermore, I anticipate that performance of these devices will decay over time due to spontaneous errors in replication of the genetic information encoding the systems. As an engineer, I would like to be able to design systems with behavior that is predictable in the face of mutation and selection. I will explore mechanisms for increasing or decreasing the susceptibility of engineered biological systems to loss of function as a result of mutation.

2 Overall Objective & Specific Aims

"There are only two ways we know of to make extremely complicated things. One is by engineering, and the other is evolution." -Danny Hillis

There is a third approach for construction of complicated things that combines forward engineering and evolution. This hybrid approach requires that the substrate be both designable and evolvable. The construction of engineered biological systems may be amenable to such an approach. I propose to develop a method for constructing engineered biological systems that combines elements of forward engineering with evolution in order to increase the likelihood of producing functional systems. Furthermore, I anticipate that system performance will become unpredictable over time due to evolutionary instability. I propose to study mechanisms for modulating the evolutionary stability of engineered biological systems, both for designing synthetic systems as well as to gain a better understanding of natural systems. Specifically, I plan to:

1) Develop tools and techniques for integrating directed evolution into the construction of engineered biological systems.
I will develop methods to generate and maintain libraries of genetic parts and devices. I will develop a generalized method to screen these libraries for desired input/output functions.

2) Construct a set of ring oscillators and bi-stable switches to validate the hybrid forward engineering/directed evolution construction approach.
I will generate libraries of four genetic inverters and screen for mutants with matching transfer functions. I will combine the screened inverters to form eight unique ring oscillators and six unique bi-stable switches.

3) *Determine system design principles for modulating the evolutionary stability of engineered biological systems.*
I will study how DNA sequence and system design can be used to tune the evolutionary stability of our engineered systems. Specifically, I will evaluate (a) sequence design that renders DNA more or less sensitive to mutation, and (b) sequence design that renders part, device, and systems performance more or less sensitive to mutation, and (c) device and system designs that render system performance more or less sensitive to mutation.

3 Background and Significance

3.1 Challenges to routine engineering of biological systems

The primary goal of biological engineering is to make routine the design and construction of reliable biological systems that behave as predicted. Engineered biological systems might manipulate information, produce materials, process chemicals, provide energy, and help maintain or enhance human health and our environment. A number of very simple genetic systems have been engineered to date, with mixed results. These systems include: inverters, bi-stable switches, oscillators, cell-to-cell communication networks, a cell density regulation circuit, and a programmed pattern forming system (Elowitz and Leibler, 2000, Gardner et al., 2000, Weiss, 2001, You et al., 2004, Basu et al., 2005).

The construction and performance of these systems demonstrate the difficulty of engineering inside living cells. A single ring oscillator took approximately two years to design and construct (Elowitz and Leibler, 2000). Construction of a bi-stable switch and matching the signals for a pair of inverters took nearly as long (Weiss, 2001, Gardner et al., 2000). Rapid construction via a combinatorial approach generated 125 different transcriptional logic circuits, yet only a fraction performed as predicted (Guet et al., 2002). Consideration of these results suggests that there are a number of challenges that prevent the routine construction of engineered biological systems. The most important of these are:

(1) Tedious and unreliable construction
(2) Overwhelming biological complexity
(3) Lack of a common signal carrier between devices
(4) Tedious characterization of device performance
(5) Unmatched levels and timing of signals between devices
(6) Evolutionary stability of the genetic information encoding the systems
(7) Variation in system performance due to stochastic biochemical reactions

The first five of these challenges can be addressed by improvements in the engineering framework utilized for design and construction of biological systems. The latter two issues reflect some of the "performance limits" encountered inside living cells; solving these challenges will require improved device and system designs or wholesale cell re-engineering. I propose to address the first five challenges by extending the BioBricks construction framework to include directed evolution. I will address the sixth challenge, evolutionary stability, by exploring DNA and system design strategies. I will not directly address the seventh issue, spontaneous variation in system performance; however some strategies for increasing evolutionary performance stability should be directly applicable to reducing the effect of short time scale physical "noise."

3.2 Current approach for engineering of biological systems

The BioBricks construction framework and its associated abstraction hierarchy were developed to meet the challenges of tedious construction, management of complexity, and a common signal carrier in engineered biological systems. Genetic elements conforming to the BioBricks standard contain specific 5' and 3' ends that

enable easy combination with other BioBrick parts via a standardized assembly procedure (Knight, 2002). The abstraction hierarchy addresses the issue of complexity by defining abstraction levels and specifying precise interfaces between those levels. This reduces the complexity seen at each abstraction level by hiding information not essential at that level in the hierarchy (Endy, 2005). Additionally, the abstraction hierarchy specifies a common signal carrier, polymerases per second (PoPS) for gene expression devices. As a result of the common signal carrier, any PoPS-based device can be connected to any other PoPS-based device (Figure 1). Taken together, the ideas of standard biological parts (i.e., BioBricks) and an abstraction hierarchy (Figure 2A) have enabled many individuals to quickly design and build a large number of engineered genetic systems. However, only a small number of the resulting systems function as desired (http://parts.mit.edu).

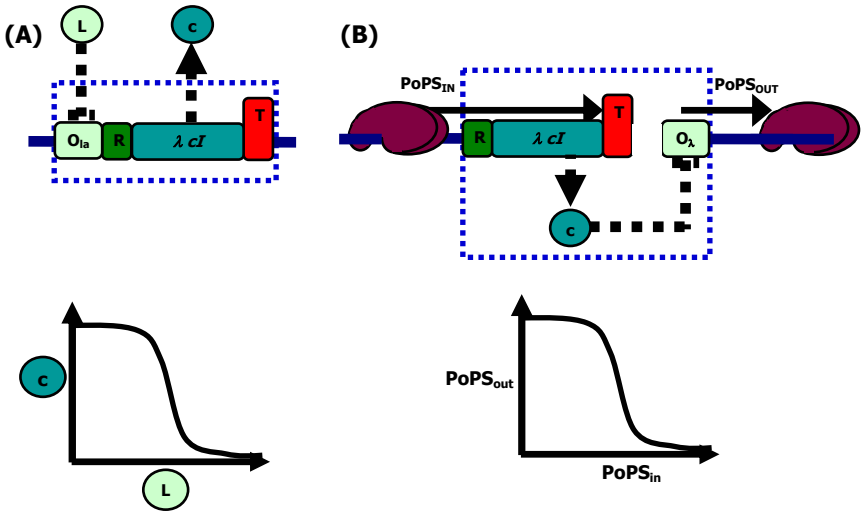

Figure 1: Two definitions for genetically encoded inverter devices. (A) A 'classical' genetically encoded inverter that takes as input the concentration of repressor protein, LacI. In the presence of LacI, expression of the downstream cI gene is inhibited. In the absence of LacI, expression of the downstream cI occurs via transcription initating at O_{Lac}, producing a high cI output signal. (B) A PoPS-based inverter. When the input PoPS-level is high, cI is produced. cI acts at O_{lambda} to keep the output PoPS-level low. The molecule-specific details of a PoPS based inverter are internal to the device and can be hidden; PoPS-based devices can thus be used in combination with (i.e., connected to) any other PoPS-based devices [Endy, iGEM 2005 Supplement, http://web.mit.edu/endy/www/igem/iGEM.supplement.pdf].

3.3 Limitations of the current approach for engineering biological systems

Figure 2(B) highlights some of the limitations of the current forward engineering approach to construction of engineered biological systems. In particular, the current approach requires precise, predictive models for performance at each level in the abstraction hierarchy. Predictive models enable engineers working at the device and part level to construct genetic elements with the specific characteristics defined at higher levels. Predictive models also enable the system level engineer to ensure that assembly of devices that meet specifications will result in a functional overall system.

The example shown in figure 2(B) depicts just one scenario where this approach fails. In this case, an adequate device-level model for an inverter is available, but a part-

level model for repressors is unavailable. Repressors that meet the particular characteristics defined at the device level are not present in an existing part collection, and there is no mechanism to generate functional repressors *de novo*. As a result, the system cannot be successfully constructed.

The forward engineering approach can break down at the device or system level, as well as at the part level. For example, if an inverter device model is not predictive then the device engineer cannot adequately define repressor protein characteristics to the part engineer. Similarly, if a system engineer has a poor model for a ring oscillator, then the assembly of inverters that meet exact specifications will still not result in a functional oscillator. Note that a deficiency in the model at *any* level of the abstraction hierarchy will prevent construction of a functional system.

In practice very few systems have performed as predicted by our models (IAP Synthetic Biology Class 2003/2004, http://rosalind.csail.mit.edu/projects/index.cgi). Additionally, parts or devices that meet desired specifications are typically not present in the existing parts collections (e.g., the MIT Registry of Standard Biological Parts). This is not surprising as most genetic elements in the Registry are derived from disparate natural systems and should not be expected to interact perfectly "out of the box."

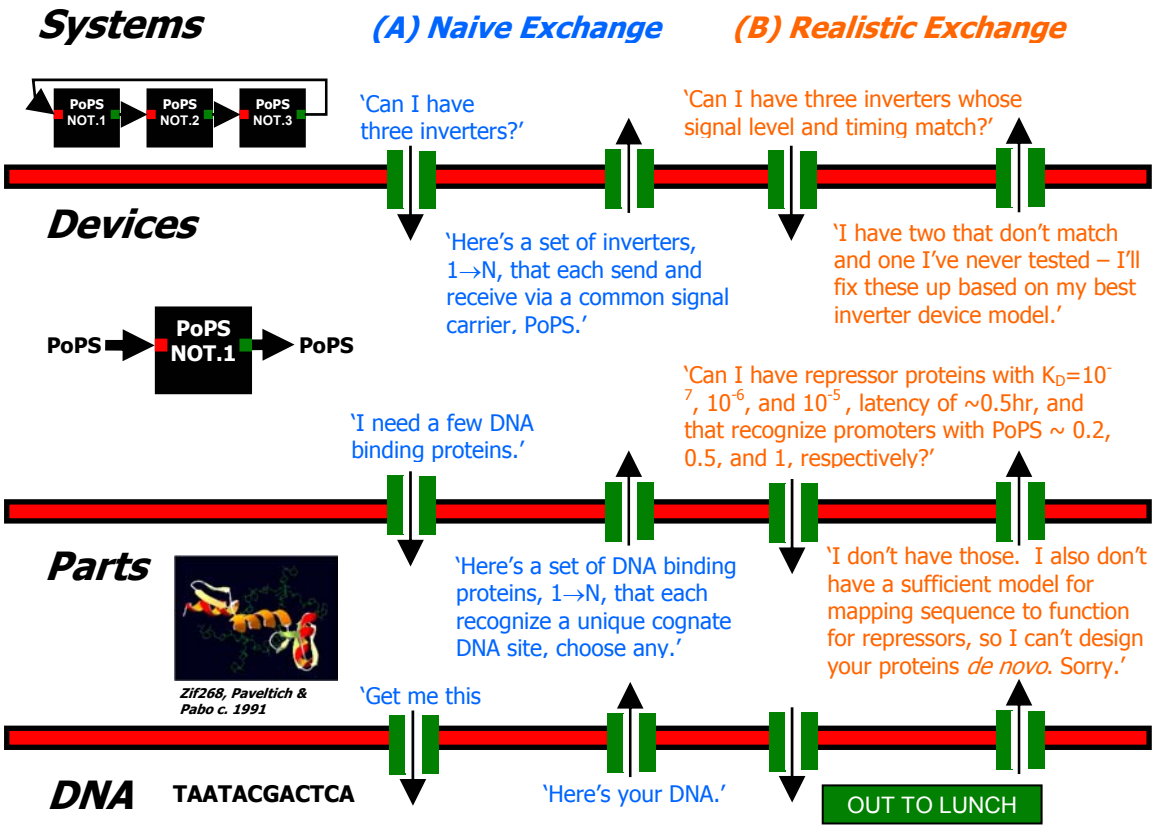

Figure 2: An abstraction hierarchy that supports construction of engineered biological systems. Abstraction levels are listed (DNA, Parts, Devices, Systems). Abstraction barriers (red) block all exchange of information between levels. Interfaces (green) enable the limited and principled exchange of information between levels. A naive exchange (blue) demonstrates successful exchanges at each barrier. A more realistic exchange (orange) depicts one example of construction failure: inadequate part level models. [Endy, iGEM 2005 Supplement, http://web.mit.edu/endy/www/igem/iGEM.supplement.pdf].

3.4 Improved construction via introduction of directed evolution

Combining directed evolution with forward engineering will address the issue of unreliable models in the exclusively forward engineering approach. Directed evolution has been applied successfully in several fields, including enzyme design, recombinant protein expression, and metabolic engineering (Kuchner and Arnold, 1997, Lin et al., 2000, Jurgens et al., 2000). In particular, library generation and screening have proven useful in areas where it is difficult to adequately model the effect of changes in system architecture on performance. The capability of directed evolution to overcome unreliable models is a clear reason for incorporating it in the construction process (Hasty, 2002).

There are limitations to an evolutionary approach as well, in particular the need for a rapid screen or selection for phenotype. This problem is compounded by the complex phenotype of our engineered devices. Fortunately, in the hybrid approach described here, directed evolution benefits from its association with forward design as well as vice versa. Some characteristics of the BioBrick construction framework are actually well-suited to enable directed evolution. For example, PoPS provides a common interface between devices and enables a single screening system to be applied to any PoPS-based device. As a result, PoPS helps avoid the intractable scenario of designing a unique screen for every device. Also, the standard BioBrick 5' and 3' ends enable easy manipulation and assembly of part, device, or system libraries. Lastly, the abstraction hierarchy provides a framework for generating diversity and screening at different levels of complexity.

One question of particular concern is at what level in the abstraction hierarchy libraries should be generated and screened. Figure 3 outlines the four possible options for combining directed evolution and forward engineering within the framework provided by the abstraction hierarchy. Figure 3 also lists specific definitions for parts, devices, and systems. Note that systems are defined as unable to be easily screened and, as a result, directed evolution cannot extend to the system level. The first option outlined in the figure is the pure forward engineering approach that is currently employed and that assumes predictive models at every level of the hierarchy (Figure 3A). The second option generates part libraries and screens to a single part with characteristics defined by a predictive device model (Figure 3B). The third option generates part libraries and screens based on predictive device models to obtain smaller libraries with desired ranges of functionality, hereafter called sub-libraries (Figure3C). These part sub-libraries are combined to generate a device library that is then screened to a single device. To be clear, by assembling a device from screened part libraries I expect to enrich the device library for functional mutants. The final option generates a device library directly at the device level and screens to a single device (Figure 3D). Note that all options assume a predictive model at the system level by definition.

Here, I will attempt to rapidly and reliably construct systems using the fourth approach. This approach was chosen based on previous work that demonstrates a successful instance of screening at the device level. For example, Yokobayashi *et al.* successfully matched a pair of initially unmatched inverters by generating a library of one inverter, combining it with the second inverter, and screening for proper function (Yokobayashi et al., 2002). If this approach does not work, I will evaluate if the other options can enable the more efficient construction of engineered biological systems.

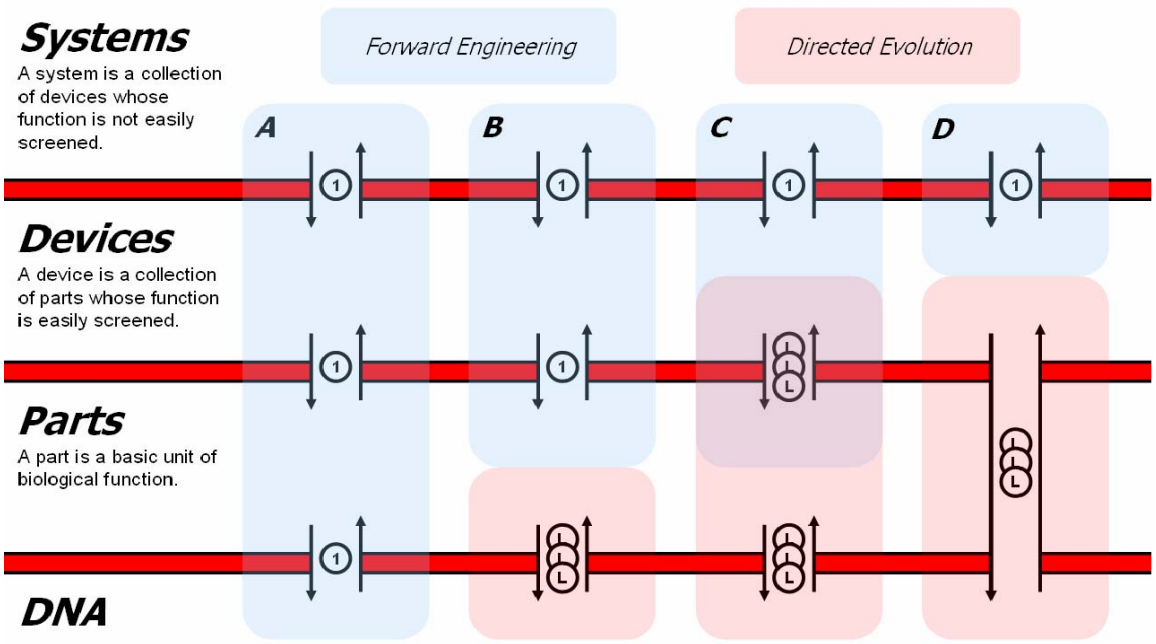

Figure 3: Four options for combining forward engineering and directed evolution in the framework of the abstraction hierarchy. (1) indicates a single instance of DNA, a part, or a device being passed across an abstraction barrier; (L) indicates a library being passed. Screens are located at the interfaces between forward engineering and directed evolution, and can be used to isolate a single mutant or a library of mutants depending on the request at the interface. (i.e. in option C libraries of parts are screened to collect smaller sub-libraries that are passed to the device level and assembled to form a device library. In option B libraries of parts are screened to isolate single mutants that are passed to the device level and assembled to form a single device.) Note the definitions for parts, devices, and systems have been specified in support of incorporating directed evolution. The directed evolution approach cannot extend to the system level by definition since systems are not easily screened.

3.5 Anticipating problems with evolutionary stability of systems

Our engineered biological systems exist inside replicating machines (i.e., living cells). Machine replication results in spontaneous errors in the genetic information encoding our systems. It is expected that these changes will likely have deleterious effects on system performance. Specifically, we have observed instability in systems containing a strong promoter, such as R0052, the bacteriophage 434 right operator (Caitlin Conboy, personal communication). Deletion mutations that removed the promoter sequence occurred in a small number of generations (a single 5ml overnight culture). Evolutionary instability in engineered systems has been observed by others as well, for example a cell density regulation circuit lost function in ~70hrs in a 50ml chemostat culture (You et al., 2004).

I anticipate that evolutionary stability will become an increasingly significant issue as our systems increase in complexity and are operated for longer periods of time. As an engineer, I would like to be able to design systems such that performance in the face of mutation and selection is predictable. Further, I would like to either decrease or *increase* the susceptibility of the system to loss of function by mutation (e.g., if I wanted a system to function only for a few generations).

To begin to solve this problem, I define two aspects of evolutionary stability: (1) genetic stability and (2) performance stability. Genetic stability is the stability of the

information encoding the system - the stability of the DNA sequence itself. Performance stability is the capacity of the system to continue to function reliably given changes in the underlying DNA sequence. Note that a system that is genetically stable would by default also have performance stability, but not vice-versa.

Evolutionary stability is a problem that natural biological systems have presumably solved, as natural systems appear to maintain many pathways that are either infrequently used, or could be considered deleterious to the individual cell (e.g., apoptosis). The methods that cells use to maintain seemingly deleterious systems are not well understood. Simple engineered biological networks offer an excellent opportunity to study the evolutionary stability of various system architectures, and to explore techniques for modulating evolutionary stability.

3.6 Mechanisms of spontaneous mutation in E. coli

The genetic stability of engineered biological systems can be modulated by increasing or decreasing the susceptibility of the system to host mutational mechanisms. In *E. coli*, spontaneous mutations occur at a rate of about $6x10^{-10}$ per base pair per doubling and arise by a wide variety of mechanisms (Drake, 1991). The mechanisms can be grouped into two categories: (1) point mutations and frameshifts and (2) gross DNA rearrangements. Point mutations and frameshifts frequently occur as results of errors in DNA replication, specifically polymerase base selectivity, mismatch repair, and proofreading (Schaaper, 1993). Gross DNA rearrangements include deletions, inversions, duplications, and transpositions. These rearrangements are typically associated with direct or inverted repeats and are thought to involve recombination mechanisms (Balbinder, 1993, Whoriskey et al., 1991, Schofield et al., 1992).

There is little quantitative data describing the relative frequencies of the various mutation types. The most well studied example is mutations leading to loss of function of the *lacI* gene. In four separate studies, a total of about 1500 *lacI* $^{-}$ mutants were sequenced and their mutations were characterized (Schaaper and Dunn, 1991, Schaaper et al., 1986, Halliday and Glickman, 1991, Farabaugh et al., 1978). About 75% of the mutants gained or lost a particular sequence, TGGC, that is repeated in several positions in the gene. The remaining mutations were divided between deletions (10%), duplications (3%), insertions (<1%), and point mutations (12%). The point mutations were further characterized by specific base pair change, with the C:G to T:A transition found to be the most frequent. This is potentially not surprising as more recent work has shown that promoting cytosine deamination (e.g., C to T mutation) in the non-transcribed strand is a general property of transcription in E. coli and is dependent on the length of time the transcription bubble stays open during elongation (Beletskii et al., 2000).

It should be noted that these are far from all of the mechanisms of mutagenesis in *E. coli*. In particular, many alternate mechanisms drive mutagenesis during SOS response to DNA damage, starvation-induced stationary phase, and in the presence of exogenous mutagens (Purmal et al., 1994, Hastings et al., 2004).

4 Research Design & Methods

4.1 Tools for library-based construction of biological systems

I will develop tools and techniques to enable the incorporation of directed evolution in the construction of engineered biological systems. I will (1) evaluate methods to generate, maintain, and characterize libraries and (2) build a screening plasmid that enables high-throughput screening of any PoPS-based device.

4.1.1 Library generation, maintenance, and characterization

Generation

I will generate device libraries by mutagenic PCR of BioBrick devices with a polymerase mix ensuring unbiased nucleotide mutations (http://www.stratagene.com, GeneMorph II Random Mutagenesis Kit). Small devices and parts (<100bp) may be generated via de novo synthesis with specifically targeted mutations. I expect that near-term advances in de novo synthesis technology will enable highly targeted mutation in larger, more complicated parts or devices.

Maintenance

I hope to maintain libraries in culture form so that the loss of library diversity accompanying plating can be avoided, however mutants must have similar growth rates so that a particular mutant does not dominate the culture in a small number of generations. Early results suggest that growth rate may be comparable for a wide range of systems that are constructed and maintained on low copy number plasmids.

Characterization

Characterization of sub-library diversity will highlight the sequence positions that are the most influential in determining device function. Individual clones can be sequenced in a high-throughput sequencing method to identify mutation "hotspots". Sub-library sequence data will prove useful for directing future mutations to particular regions of the sequence in order to increase the diversity of device libraries. Sub-library sequencing will also provide empirical data for improving performance models.

4.1.2 Build and test the PoPS screening plasmid

The PoPS screening plasmid will enable the screening of device libraries for particular input/output functions. (Fig4) The system will consist of 3 main components:
(1) a controllable, linear PoPS input generator
(2) an internal BioBrick cloning site for easy insertion and removal of device libraries
(3) PoPS input and output measurement devices

The plasmid itself will be based on pSB2K3 (Registry, http://parts.mit.edu). pSB2K3 is derived from the variable copy plasmid system, pSCANS. The plasmid contains the F′ replication origin (copy number ~2) and also the P1 lytic origin (copy number >100). Replication at the lytic origin can be induced by IPTG. This will facilitate screening of devices at low copy number (expected operating conditions for our systems) while allowing for induction to high copy number to increase DNA preparation yield for subsequent construction steps. A similar dual-fluorescence reporter system has been constructed by others and used to successfully characterize and isolate insert-

bearing clones following subcloning of a BAC in a genome sequencing procedure.(Choe et al., 2005)

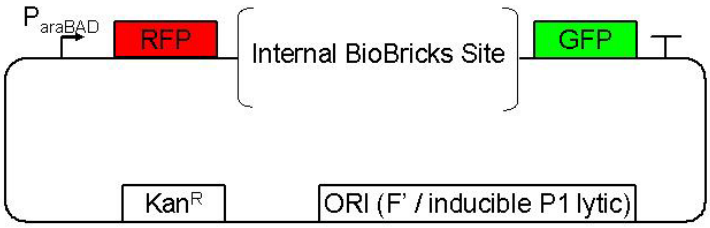

Figure 4: PoPS screening plasmid schematic.

PoPS input generator

In order to screen device libraries for particular input/output functions it is necessary to provide a range of PoPS input signals to the device. A well-controlled inducible PoPS input signal that can be modulated by an external inducer would provide optimal control. I will initially implement the P_{BAD} system developed by Keasling *et al*, and will establish experimental conditions that allow for reproducible, controlled induction (Khlebnikov et al., 2002). The P_{BAD} system relies on the naturally occurring arabinose-inducible araBAD promoter. In future versions of the screening plasmid I will move to a purely synthetic system, rather than one that relies on a sugar and a naturally occurring metabolic network. The reason for this preference is to reduce cross-talk and other unpredictable host effects associated with using a native system.

Internal BioBrick cloning site

The internal BioBricks cloning site enables any part library to be easily inserted into the screening plasmid, as well as allows sub-libraries to be easily removed for use in subsequent assembly steps. In the process of inserting the internal BioBricks cloning site in the screening system, we developed several new tools to standardize the process for future BioBricks users. The first is a BioBricks part that enables the insertion of a new BioBrick cloning site internal to a device, and the second removes BioBrick cloning sites on the end of a part or device (see http://parts.mit.edu, BBa_I13450 and BBa_I13452).

PoPS input and output measurement devices

PoPS input and output signals will be measured indirectly by expression of fluorescent proteins GFP and mRFP1, respectively. The fluorescent proteins will enable an internal control for variation in expression from the PoPS input device as well as enable high speed cell sorting based on the levels of GFP and mRFP1. In future versions of the screening plasmid I will include a selectable marker, such as antibiotic resistance, following the PoPS output signal measurement device. This marker will enable selection based on output signal which will reduce library size for subsequent screening.

Since I will be using two different fluorescent proteins as surrogates for measurements of PoPS, I will need to measure the constant of proportionality between the relative fluorescence levels of the two proteins. This will enable PoPS input signals measured in GFP fluorescence to be related to PoPS output signals measured in mRFP1 fluorescence. To measure the constant of proportionality, I will provide an identical

PoPS signal to both GFP and mRFP1, by simply building the screening plasmid with nothing inserted in the internal BioBrick cloning site. The PoPS signal measured at the input measurement device (GFP) should be identical to that received at the output measurement device (mRFP1), and the ratio of the fluorescence intensities will yield the constant of proportionality. This version of the screening plasmid has been constructed, and early results suggest the ratio of fluorescence intensities is constant over a range of PoPS levels under certain conditions (see Preliminary Results).

I will also evaluate various decisions in the screening plasmid architecture such as the strength of the RBS driving expression of the PoPS output measurement device and whether GFP or mRFP1 would be better for measuring input or output PoPS levels. Finally, I will use existing inverter devices to quickly evaluate the capability of the screening system to characterize the PoPS input/output function of a device. Specifically I intend to test BBa_Q04530 (p22 c2), BBa_Q04121 (lacI), BBa_Q04400 (tetR), and BBa_Q04510 (cI). Successful characterization of these inverters using the screening plasmid will pave the way for screening of libraries of inverter devices. Additionally, it will demonstrate a method for rapid characterization of devices, addressing one of the main challenges to construction of engineered biological systems.

4.2 Construction of ring oscillators and bi-stable switches

I will generate libraries of four genetic inverters and screen for mutants with matching transfer functions (Figure 5). The screened inverters will then be combined to form eight unique ring oscillators and six unique bi-stable switches. Ring oscillators and bi-stable switches were chosen because successful examples and plausible models of these devices exist in the literature (Elowitz and Leibler, 2000, Gardner et al., 2000). I hope to demonstrate the utility of the library-based construction approach by rapidly and reliably generating many more functional instances of these systems. It should be noted that the original systems took several years to construct and debug. If construction of these systems is unsuccessful, I will examine the practical limits on varying device transfer functions by mutation and debug the process by evaluating the best abstraction level(s) for combining evolution with forward design.

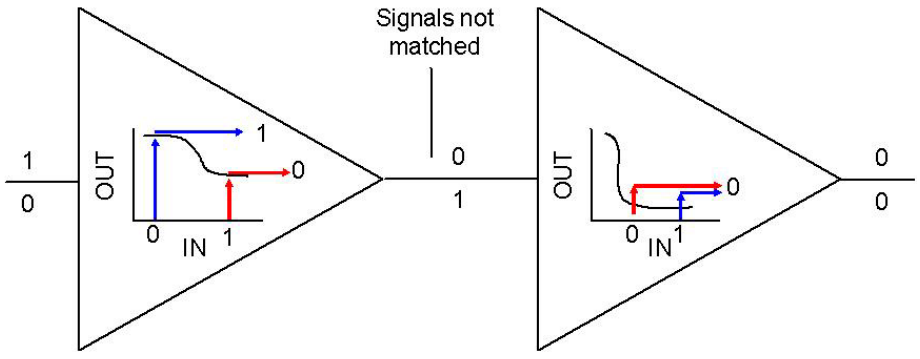

Figure 5: Example of unmatched signal levels between devices. The second inverter interprets both the high (1) and low (0) input from the first inverter as high inputs. So as a result no matter what signal enters the device the output from the second inverter is always low.

4.2.1 Inverter library generation and screening

I will generate four inverter libraries based on devices: BBa_QI9002 (penI), BBa_Q04121 (lacI), BBa_Q04400 (tetR), and BBa_Q04811 (temp-sensitive cI). I will screen the inverter libraries for mutants with matching PoPS signals based on simple models for matched inverters in series (Figure 5). Library generation and screening will take place directly at the device level, with no part level libraries (e.g. Figure 3D). This method requires fewer steps than approaches that involve assembly of parts or part libraries to generate functional devices, and previous work generating diversity at this level has been successful (Yokobayashi et al., 2002).

Each of these inverters will be mutated via mutagenic PCR with the standard BioBrick primers, VF1 and VR, in order to generate inverter device libraries. The inverter library PCR product will be digested with BioBrick enzymes and ligated into the screening plasmid. Highly competent cells will be transformed via electroporation with the ligation product. Successful transformants will be grown in supplemented M9 media and induced with a low arabinose concentration to provide a low PoPS input signal to the inverter device library. The library will then be screened by high-speed fluorescence activated cell sorters (FACS) for mutants whose GFP and mRFP1 fluorescence levels correspond to the target PoPS input/output transfer curve predicted by the matched inverters in series model. The sub-library that has been screened for the target functionality will be grown with a higher arabinose induction and again screened via FACS for another point on the transfer curve. The process of sub-library generation, induction with a new arabinose concentration, and screening for function can be repeated as many times as desired to achieve increased fidelity to the desired input/output function. Finally, individual clones will be collected from the device library and assayed for their input/output functions.

If I am unable to generate adequate functional diversity to achieve matched transfer functions, I will evaluate methods for enriching the functional diversity of the inverter libraries. For instance, generating libraries at the part level and combining screened fractions of these libraries to generate an inverter library as described in 4.2.4, below.

4.2.2 Construction of eight unique ring oscillator systems

Figure 6 depicts an example of the complete process of constructing a system via a library-based approach, specifically the construction of a single ring oscillator. Here, I will generate eight unique ring oscillators from the four matched inverter mutants constructed as described in 4.2.1. I will combine the inverters in eight unique permutations and assay for correct ring oscillator function by co-transformation with a reporter device. The reporter device will consist of a GFP coding region downstream of a promoter regulated by one of the repressors in the ring oscillator. Correct function of the ring oscillator will lead to periodic expression of GFP by cells growing on an agarose pad.(Elowitz and Leibler, 2000)

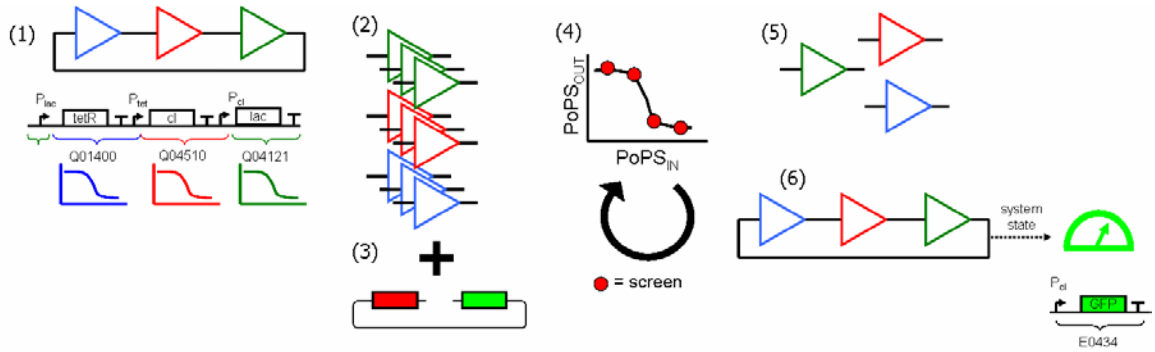

Figure 6: Overall process for library-based construction of a ring oscillator. (1) A system level model specifies the transfer curves for each of the three inverter devices composing the oscillator. (2) Libraries of each of the inverters are generated. (3) Libraries are inserted into the screening plasmid. (4) Libraries undergo rounds of screening at various points on the target transfer curve. (5) Single inverter mutants are isolated with the correct transfer functions. (6) Inverters are combined to construct the final device and it is tested for appropriate function.

4.2.3 Construction of six unique bi-stable switch devices and a library of a single switch

I will construct a set of six unique bi-stable switches by combining pairs of the four matched inverters. (e.g. lacI/tetR, lacI/cI, lacI/penI, tetR/cI, terR/ penI, cI/ penI) These devices can be assayed for function by co-transformation with a reporter device consisting of GFP downstream of one of the promoters that is targeted by the pair of inverters. I will change states in the switch by inducing with an external inducer and measuring fluorescence. The bi-stable switch has an advantage over the ring oscillator in that its function can be easily screened. I will use the simple functional screen to construct a library of a single switch (e.g. lacI/tetR). Generating libraries of switches will be particularly important if I am unable to construct functional switches by composing single inverter mutants.

I will generate a library of a single bi-stable switch by first generating libraries of two inverters. Instead of screening these inverter libraries to a single mutant, I will collect a smaller sub-library of mutants with a target range of matching transfer functions for both inverters. These two sub-libraries will be combined to make a library of bi-stable switches. This library can then be screened for functional switches, and I suspect that this method will increase the likelihood of matching the inverters and generating a functional switch. I also expect that the final switch library will be composed of switches with a variety of switching characteristics (e.g. speed, switch point) and I will assay individual mutants for their switching characteristics.

4.2.4 Debugging the library-based approach

If construction of systems is not successful via generation of diversity directly at the device level, I will attempt to debug the process by evaluating:

(1) Expected functional diversity for some common parts and devices.
(2) Enrichment for function at the device level by screening at the part level.
(3) Increasing the functional diversity at the device level via multiple rounds of mutation or increased mutation rates in library generation.

Expected diversity for some common parts and devices

Knowledge of the expected functional diversity for some common parts and devices will help to guide systems engineers in designing systems that are more likely to be successfully constructed. I will evaluate the generation of functional diversity for some common parts (promoters and terminators) as well as for a variety of inverters (lacI, tetR, cI, p22 c2).

I will generate the part libraries via *de novo* synthesis with mutations in suspected functional regions based on qualitative part models. For instance, qualitative knowledge such as the position of the -10 and -35 regions or the binding site for tet repressor would enable targeted mutagenesis in BBa_R0040 (tet promoter) to effect promoter strength or repressor binding, respectively. The variability in PoPS signal from these libraries will help to quantify the functional diversity that can be expected from promoter and terminator parts. Previous researchers have similarly characterized part performance using libraries, in particular, Imburgio *et al* evaluated a library of all single base pair mutations from the consensus T7 promoter (Imburgio et al., 2000).

I will evaluate the functional diversity of devices by generating libraries of the following inverters via mutagenic PCR: BBa_Q04530 (p22 c2), BBa_Q04121 (lacI), BBa_Q04400 (tetR), and BBa_Q04510 (cI). A large number of individual mutants will be isolated from the libraries, sequenced, and have their transfer curves characterized. If the fraction of trivial (i.e., always ON or always OFF) mutants is too large I will screen the library for a wide range of functional transfer curves. Individual mutants will then be isolated from this functional sub-library. The variety of transfer curves exhibited should provide empirical evidence of the functional diversity expected from inverter devices.

Enrichment for function at the device level by part level screening

Enrichment for functional mutants at the device level by part level screening will increase the fraction of mutants with the correct transfer function in device level libraries. I will evaluate this approach by assembly of screened part sub-libraries. I will compare the functional diversity of a library generated by mutagenesis of BBa_Q04400 (tetR inverter) directly to that of a library generated by combining a library of the first three parts of BBa_Q04400 (tetR inverter) with a screened sub-library of the last part, BBa_R0040 (tet promoter). The BBa_R0040 (tet promoter) sub-library will have been screened for a range of PoPS output levels based on a device model that predicts the optimal promoter strength to achieve a target transfer function (Shetty, 2005). I will compare the fraction of functional mutants generated by this approach to the fraction of functional mutants in the BBa_Q04400 (tetR inverter) library generated directly at the device level. I suspect that the library generated by assembly of screened part sub-libraries will have a higher fraction of functional mutants.

Increased functional diversity by multiple rounds of directed evolution or increased mutation rate

For both parts and devices I will evaluate the effectiveness of multiple rounds of directed evolution or increased mutation rates in PCR mutagenesis as methods for increasing functional diversity in libraries. I will evaluate multiple rounds of directed evolution by screening for a small fraction of mutants that are the closest to the desired function and then mutating this fraction and repeating the screen. This process can be

repeated for many rounds of mutation until the desired function is evolved. I expect that increasing the mutation rate of error prone polymerases in mutagenic PCR will increase the phenotype space explored by the library; however it may lead to more non-functional mutants.

4.3 Control of the evolutionary stability of engineered biological systems

I will evaluate methods to control the genetic and performance stability of engineered biological systems by (1) evaluating DNA sequence design that increases or decreases the genetic and performance stability of the system and (2) evaluating device and system design that increases or decreases performance stability of the system.

4.3.1 Sequence design to modulate genetic and performance stability

Develop an algorithm for codon optimization of evolutionary stability

I will develop an algorithm to codon optimize a system for increased or decreased genetic and performance stability. Codon optimization enables the underlying DNA sequence to be changed while keeping the amino acid sequence constant, by taking advantage of redundancy in the genetic code. I will use codon optimization to modulate the evolutionary stability of the system while leaving the system performance unaffected.

By varying the DNA sequence through codon optimization I will modulate both genetic and performance stability. For instance, genetic stability will be increased or decreased by changing the number of repeat homology regions, thereby encouraging or discouraging recombination. Similarly, the frequency of G:C pairs may encourage or discourage point mutations and influence genetic stability. Performance stability can be varied by choosing more or less 'volatile' codons (Plotkin et al., 2004). Each codon can be given a volatility score based on its likelihood to mutate into a stop codon or a different amino acid, in particular ones with very different chemical properties. By choosing a set of codons that are less volatile I expect system performance to be more reliable in the face of mutations.

The algorithm for codon optimization will need to consider both genetic and performance stability. For example, choosing less volatile codons to increase performance stability could create repeated homologous regions that result in genetic instability, leading to a decrease in overall evolutionary stability.

Although the amino acid sequence remains unchanged following codon optimization, there may be second order effects from altering the codon sequence that affect system performance. For example, genes with codon frequencies that are not equivalent to their respective tRNA frequencies in the host cell may have reduced expression rates due to inefficient use of cellular tRNA resources (Kurland, 1991). In refining the algorithm I will develop a version that takes into account this 'codon bias' of the host organism in determining the optimal codon sequence. Another concern is altered mRNA secondary structure that may influence regulation of protein synthesis and mRNA stability.

Experimentally validate model predictions

I will evaluate the effectiveness of the algorithm for codon optimization by measuring the stability of different codon sequences of a simple engineered system that

are generated by *de novo* synthesis. The first version of this system will be Q04121, a lacI-based inverter, controlling expression of an antibiotic. I chose lacI because it is a well studied protein, and its partially known sequence-to-function relationship will aid in interpreting data. Additionally, the sequence features which effect genetic stability in *lacI* are better specified since it is one of the few genes whose mutational frequencies have been measured. I will select for *lacI* mutants that are unable to repress antibiotic expression by plating on media containing the antibiotic. The performance stability can be evaluated by simply counting the number of colonies that grow. Mutants will be sequenced to better evaluate the genetic stability of the device. Finally, constitutive expression of GFP may provide an alternative system where loss of GFP function by mutation can be assayed via FACS.

I will synthesize codon variants that have a range of predicted stabilities, both higher and lower than the wild-type codon sequence and experimentally measure their evolutionary stability. The experimental results will be used to evaluate as well as improve the codon optimization algorithm.

4.3.2 Device and system design to modulate performance stability

Performance stability can also be modulated by the design of the device or system itself. Additionally, better understanding of the system design principles for modulating evolutionary stability may shed some light on the mechanisms natural systems utilize for stabilizing genetic elements. I will extend and evaluate the following system designs strategies if codon optimization proves ineffective.

Redundancy

I will design a system, such as a bi-stable switch, containing redundant copies of particular device components. The performance stability of the system will presumably increase because all the copies would need to break before the system loses function. However, it is possible that the second copy of the gene will result in increased recombination in the device, requiring codon optimization of the different copies to avoid homologous DNA sequences.

Simple selection circuits

Incorporating selective pressure for device components could enable long-term performance stability. For instance the performance of a ring oscillator system could be maintained by adding a second device that selects for functionality of the repressor proteins in the system. This device will contain repressor-regulated promoters upstream of genes encoding proteins whose presence is deleterious to the cell, providing a selective pressure to maintain repressor function. The design challenge would be to design a selection device such that it does not interfere with the main system performance. In this case, the promoters regulating deleterious gene expression would need to be repressed by low levels of each repressor, and the ring oscillator would have to function reliably with some low level of each repressor protein present independent of state. Otherwise, the selection would be applied each time the oscillator changed state. This approach is not as general as codon optimization; however the inclusion of selection may enable system performance to be extended indefinitely, so long as the selective pressure remains in place.

Co-evolve an engineered device and a host cell

I will attempt to evolve a host cell to gain a competitive advantage from maintaining the function of a simple engineered device by propagating the device over a large number of generations with selective pressure for appropriate device performance. Previous work with a tetracycline resistance plasmid by Lenski *et al* suggests this approach may be feasible (Lenski et al., 1994). After five hundred generations in media containing tetracycline the host cell evolved a competitive advantage from the tetracycline resistance plasmid independent of the antibiotic being present. This was shown by competition experiments between the plasmid-containing cell and an isogenic plasmid-free counterpart in media without tetracycline.

I will attempt to evolve a host cell to stabilize an inverter that regulates expression of a selectable marker that also has a counter-selection. I will induce the inverter to switch between ON and OFF states and select for expression or non-expression of the marker. This selection for function will be repeated periodically while the cells are grown over a large number of generations. Finally, the growth of cells containing the device will be compared to isogenic device-free counterparts in non-selective conditions to evaluate if the host has evolved to gain a competitive advantage from maintaining the device. In support of providing more sophisticated selections to devices in continuous culture, I will design and operate a microfluidic chemostat integrated with a cell sorter (i.e. a Sort-o-stat, see Appendix). This microscope-based system will enable me to test whether or not more complicated screens and selections will be of practical use in service of helping to evolve engineered biological systems.

Mechanisms suggested by natural systems

Previous work suggests some natural mechanisms for increasing the evolutionary stability of biological systems. For example, the demand theory of gene expression developed by Savagneau explores why a single gene might be regulated by a transcriptional activator or repressor (Savageau, 1983). Demand theory outlines the system architecture that best leverages selective pressure on protein expression towards maintenance of the control system for protein expression. Although demand theory has been supported by examining native regulatory systems, I will provide a more direct proof by demonstrating demand theory via experimental evolution of a simple engineered system.

Additionally, a cue can be taken from the many viral genomes that utilize multiple reading frames for overlapping genetic elements in the same sequence. This mechanism is thought to serve as a means of information compression due to selection for small genome size in viral replication. It may also play a role in stabilizing non-essential genetic elements via 'interlacing' with essential elements. For instance, a non-essential gene that shared sequence with an essential gene may gain resistance to frameshift mutations since such a mutation would typically result in a loss of function of the essential gene, selecting against the mutant. This approach will be used to stabilize engineered systems by interlacing system components with essential genes or selectable markers.

5 Preliminary Studies

5.1 Construction and testing of PoPS screening plasmid

5.1.1 Materials and methods

Systems

I have constructed two test version of the PoPS screening plasmid, BBa_I13513 and BBa_I13514. The systems are similar to that depicted in figure 4; BBa_I13513 has RFP preceding GFP and BBa_I13514 has the order of the fluorescent proteins reversed. In both systems fluorescent protein expression is regulated by the P_{BAD} promoter. The systems were constructed in pSB1A3 (Amp^R, pMB1 *ori*, 100-300 copies per cell). I am currently transferring the PoPS screening system to two lower copy plasmids, pSB4A2 (Amp^R, pSC101 *ori*, 18-22 copies per cell) and pSB2K3 (Kan^R, F′ *ori*, ~2 copies per cell and an inducible P1 lytic *ori*, >100 copies per cell).

I obtained the pJat18 plasmid from the laboratory of Jay Keasling. pJat18 contains *araE*, an arabinose permease, under control of a constitutive promoter, P_{CP18}, and is a pJN105-derived plasmid (Gm^R, Erm^R, pBBR-1 *ori*, low copy number). This plasmid is part of the P_{BAD} expression system developed by Keasling *et al* to avoid the native positive feedback mechanism that leads to all-or-none expression from the P_{BAD} promoter. Transformation of pJat18 into an $araE^-$ strain enables consistent induction with arabinose across all cells in the population (Khlebnikov et al., 2002).

Strains

E. coli strain CW2553 (*araE201* Δ*araFGH*::*kan*) is a K-12 derivative with the wild-type genes for arabinose transport knocked out or mutated. This strain was used for all the experiments that follow.

Experimental Conditions

I co-transformed CW2553 with pJat18 and BBa_I13513 or BBa_I13514, both on pSB1A3, in order to enable induction of the fluorescent proteins with arabinose. Cells were grown overnight at 37C in 5ml of supplemented M9 minimal media (0.1% cas amino acids, 0.1% thiamine, and 0.4% glycerol) and antibiotics ampicillin, kanamycin, and gentamycin. I diluted the cultures back 1 in 10 into supplemented M9 media with 0.2% glucose rather than glycerol to strongly repress expression from the P_{BAD} promoter. After 3 hours, the cells were collected by centrifugation and resuspended to an OD of 0.01 in fresh supplemented M9 with antibiotics and arabinose concentrations of 0%, 0.00003%, 0.0001%, 0.0003%, 0.001%, 0.003%, 0.01%, 0.03%, and 0.1% w/w. Cells were placed at 37C and samples were collected at 5, 6, and 7 hours following induction with arabinose. GFP and mRFP1 fluorescence levels for individual cells were measured on a Beckton Dickinson FACScan flow cytometer. [http://www.bd.com] Cells were excited with a 488nm laser and emission was collected though a GFP filter (530nm/30) and an RFP filter (650nm/LP). Fluorescence intensities were calibrated against beads with known intensities to account for day-to-day machine variation. [http://www.spherotech.com].

5.1.2 Results

Induction via the P_{BAD} expression system

Induction of GFP and mRFP1 from BBa_I13513 and BBa_I13514 was shown to be inducible for a range of arabinose concentrations from 0-0.1% w/w. The results matched well with previous work by Keasling *et al*. Figure 7 (A&B) shows a 10-fold increase in protein expression for arabinose induction below 0.03%, and a decline in expression levels at higher arabinose concentrations similar to results observed in previous work (Khlebnikov et al., 2002). Figure 7(C&D) show that the all-or-none expression phenomenon present in wild-type arabinose induction is not present in the *araE* deleted, pJat18-containing strain. These results suggest that the P_{BAD} induction system has been implemented successfully, and is capable of delivering a well-controlled, 10-fold range of PoPS signals to devices in the screening plasmid.

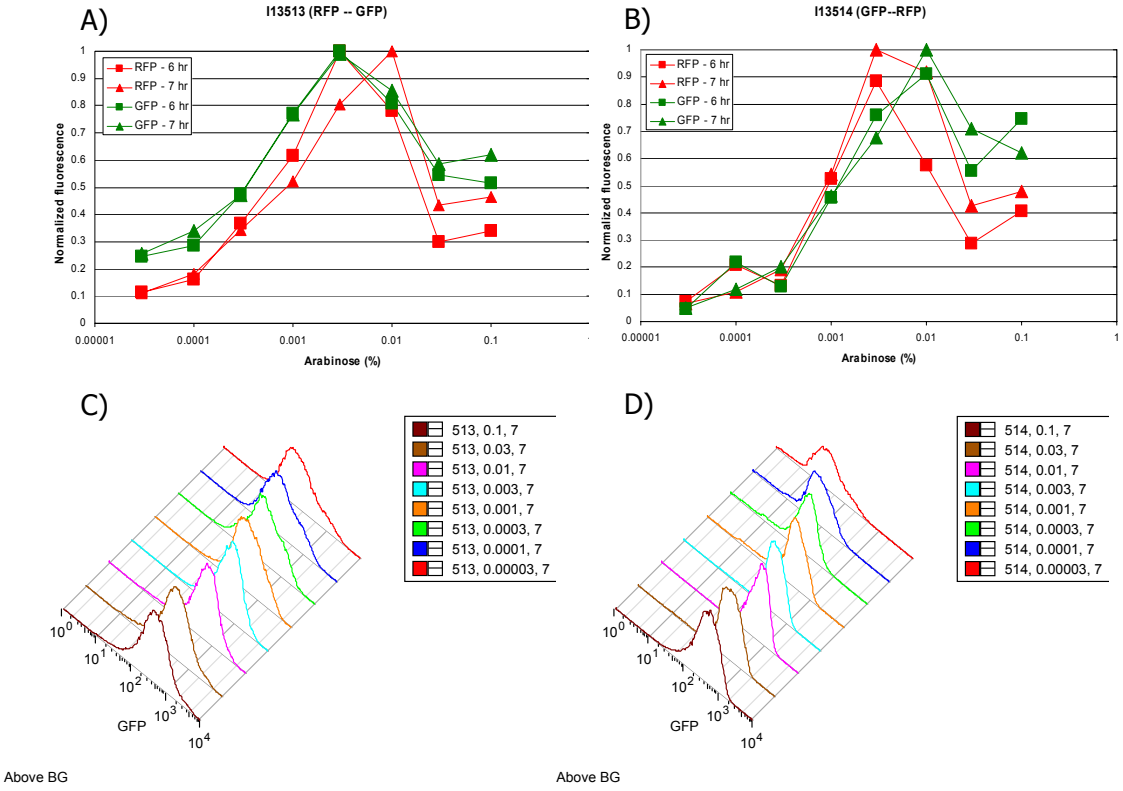

Figure 7: (A) I13513 and (B) I13514 were induced with a range of arabinose concentrations and samples were collected 6 (squares) or 7(triangles) hours after induction. Values shown are mean GFP and mRFP1 fluorescence levels normalized to the highest mean GFP or mRFP1 fluorescence level for their respective time point. (C&D) The single cell fluorescence histograms plotted for one time point, 7 hours post induction, for (C) I13513 and (D) I13514 confirms that induction is consistent across the population of cells since a single peak is present, rather than the bimodal distribution characteristic of all-or-none induction

Evaluate constant of proportionality between fluorescence levels of GFP and mRFP1

Induction of GFP and mRFP1 from BBa_I13513 and BBa_I13514 provides a preliminary measurement of the constant of proportionality between the fluorescent levels of the fluorescent proteins in response to identical PoPS signals. This ratio is expected to remain constant across a range of PoPS signals, however it is likely to be dependent on the order of the fluorescent proteins (e.g. it should be different for BBa_I13513 and BBa_I13514). The two arrangements were constructed simply to evaluate whether one order would provide a more consistent ratio of GFP to mRFP1.

The ratio of GFP/mRFP1 was found to be relatively well conserved in BBa_I13513 with a value of approximately 13 over a 10-fold range of PoPS input signals, as well as across several time points following induction (Figure 8). The results were not as promising for BBa_I13514, which displayed a greater variability in the GFP/RFP ratio across the same range of arabinose concentrations (Figure 8).

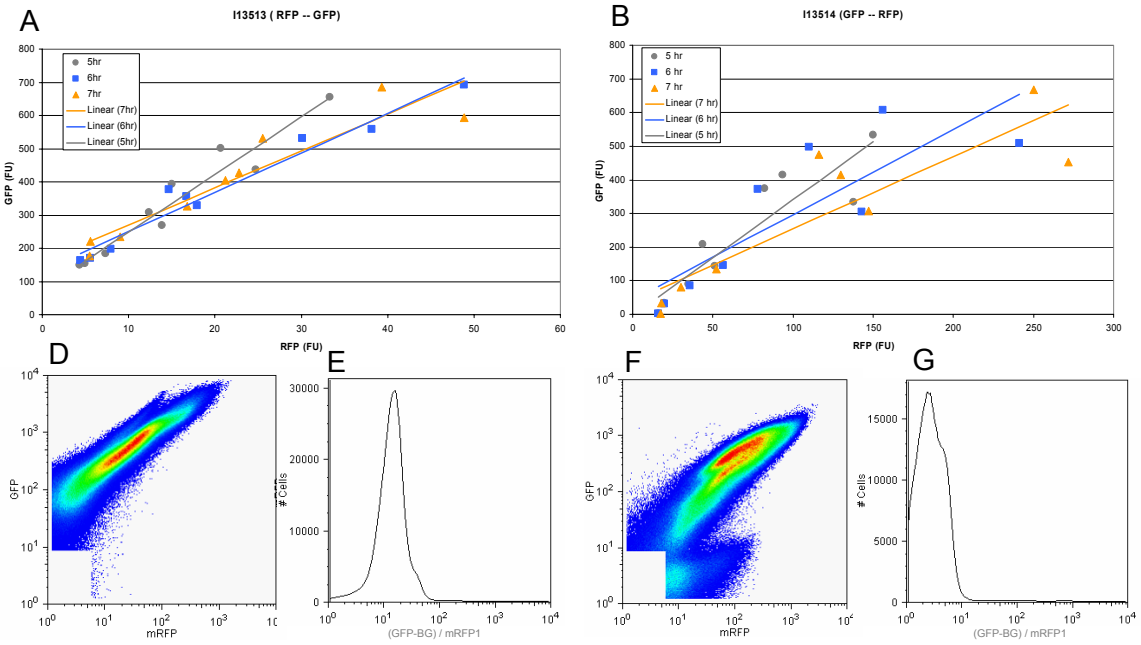

Figure 8: Figure 8 (A & B) shows the geometric mean of the GFP and mRFP1 fluorescence intensities in relative fluorescence units across three time points and six arabinose concentrations. The lines shown are the linear best fit for each time point. The dot plots for BBa_I13513(D) and BBa_I13514(F) display the complete (all arabinose concentrations and times) single cell data set concatenated on a single plot. The histograms show the ratio of GFP/RFP corrected for background GFP for BBa_I13513 (E) and BBa_I13514 (G) for the complete data set.

It is not known at this time why BBa_13514 has a less consistent ratio than that of BBa_I13513, however this data is from a single experimental run and it may simply be a result of experimental error. I will be evaluating if the same effect is seen at lower copy and after standardizing the mRNA stability of the devices (see note below). Furthermore, I will measure the approach to steady-state fluorescent protein levels by growing cultures and measuring population fluorescence readings in a 96-well Victor3 fluorometer [http://www.perkinelmer.com]. Information about the time until steady-state will help determine the most effective time to collect samples post-induction to achieve reliable measurements.

Note on matching the mRNA stability of reporter protein coding mRNA

mRNA stability is a particularly important concern in the screening system since it is critical that the relationship between fluorescence intensity and PoPS remains constant independent of the device that is being screened. For instance, in the current screening plasmid design if a device does not end in a terminator or promoter the second fluorescent protein will be part of a polycistronic message including a component of the device. This may influence the stability of the mRNA encoding the fluorescent protein and as a result may affect the relationship between $PoPS_{OUT}$ of the screened device and the level of fluorescent protein used as a surrogate for this signal. Similarly, if the first component of the device is not a terminator then the first fluorescent protein will be a part of a polycistonic message including a component of the device, and face similar issues.

I will address this issue by including an RnaseE site downstream of the first fluorescent protein and a second RnaseE site upstream of the second fluorescent protein, on either side of the BioBrick multi-cloning site. Additionally, stem-loop structures will be added adjacent to the RnaseE sites to further stabilize the mRNA following cleavage of the polycistronic mRNA. The addition of new RnaseE sites and stem-loop secondary structures has previously been shown to stabilize the expression levels of a pair of reporter proteins (Smolke and Keasling, 2002).

6 Literature cited

Balagadde, F. K., L. You, et al. (2005). "Long-term monitoring of bacteria undergoing programmed population control in a microchemostat." Science 309(5731): 137-40.

Balbinder, E. (1993). "Multiple pathways of deletion formation in Escherichia coli." Mutat Res 299(3-4): 193-209.

Basu, S., Y. Gerchman, et al. (2005). "A synthetic multicellular system for programmed pattern formation." Nature 434(7037): 1130-4.

Beletskii, A., A. Grigoriev, et al. (2000). "Mutations induced by bacteriophage T7 RNA polymerase and their effects on the composition of the T7 genome." J Mol Biol 300(5): 1057-65.

Choe, J., H. H. Guo, et al. (2005). "A dual-fluorescence reporter system for high-throughput clone characterization and selection by cell sorting." Nucleic Acids Res 33(5): e49.

Drake, J. W. (1991). "A constant rate of spontaneous mutation in DNA-based microbes." Proc Natl Acad Sci U S A 88(16): 7160-4.

Elowitz, M. B. and S. Leibler (2000). "A synthetic oscillatory network of transcriptional regulators." Nature 403(6767): 335-8.

Endy, D. (2005). "iGEM 2005 Supplement, http://web.mit.edu/endy/www/igem/iGEM.supplement.pdf."

Farabaugh, P. J., U. Schmeissner, et al. (1978). "Genetic studies of the lac repressor. VII. On the molecular nature of spontaneous hotspots in the lacI gene of Escherichia coli." J Mol Biol 126(4): 847-57.

Gardner, T. S., C. R. Cantor, et al. (2000). "Construction of a genetic toggle switch in Escherichia coli." Nature 403(6767): 339-42.

Guet, C. C., M. B. Elowitz, et al. (2002). "Combinatorial synthesis of genetic networks." Science 296(5572): 1466-70.

Halliday, J. A. and B. W. Glickman (1991). "Mechanisms of spontaneous mutation in DNA repair-proficient Escherichia coli." Mutat Res 250(1-2): 55-71.

Hastings, P. J., A. Slack, et al. (2004). "Adaptive amplification and point mutation are independent mechanisms: evidence for various stress-inducible mutation mechanisms." PLoS Biol 2(12): e399.

Hasty, J. (2002). "Design then mutate." Proc Natl Acad Sci U S A 99(26): 16516-8.

Imburgio, D., M. Rong, et al. (2000). "Studies of promoter recognition and start site selection by T7 RNA polymerase using a comprehensive collection of promoter variants." Biochemistry 39(34): 10419-30.

Jurgens, C., A. Strom, et al. (2000). "Directed evolution of a (beta alpha)8-barrel enzyme to catalyze related reactions in two different metabolic pathways." Proc Natl Acad Sci U S A 97(18): 9925-30.

Khlebnikov, A., T. Skaug, et al. (2002). "Modulation of gene expression from the arabinose-inducible araBAD promoter." J Ind Microbiol Biotechnol 29(1): 34-7.
Knight, T. (2002). "Idempotent Vector Design for Standard Assembly of Biobicks."

Kuchner, O. and F. H. Arnold (1997). "Directed evolution of enzyme catalysts." Trends Biotechnol 15(12): 523-30.

Kurland, C. G. (1991). "Codon bias and gene expression." FEBS Lett 285(2): 165-9.

Lenski, R. E., S. C. Simpson, et al. (1994). "Genetic analysis of a plasmid-encoded, host genotype-specific enhancement of bacterial fitness." J Bacteriol 176(11): 3140-7.

Lin, Z., T. Thorsen, et al. (2000). "Functional expression of horseradish peroxidase in E. coli By directed evolution." Biotechnol Prog 16(4): 680.

Plotkin, J. B., J. Dushoff, et al. (2004). "Detecting selection using a single genome sequence of M. tuberculosis and P. falciparum." Nature 428(6986): 942-5.

Purmal, A. A., Y. W. Kow, et al. (1994). "Major oxidative products of cytosine, 5-hydroxycytosine and 5-hydroxyuracil, exhibit sequence context-dependent mispairing in vitro." Nucleic Acids Res 22(1): 72-8.

Savageau, M. A. (1983). "Regulation of differentiated cell-specific functions." Proc Natl Acad Sci U S A 80(5): 1411-5.

Schaaper, R. M. (1993). "Base selection, proofreading, and mismatch repair during DNA replication in Escherichia coli." J Biol Chem 268(32): 23762-5.

Schaaper, R. M., B. N. Danforth, et al. (1986). "Mechanisms of spontaneous mutagenesis: an analysis of the spectrum of spontaneous mutation in the Escherichia coli lacI gene." J Mol Biol 189(2): 273-84.

Schaaper, R. M. and R. L. Dunn (1991). "Spontaneous mutation in the Escherichia coli lacI gene." Genetics 129(2): 317-26.

Schofield, M. A., R. Agbunag, et al. (1992). "DNA inversions between short inverted repeats in Escherichia coli." Genetics 132(2): 295-302.

Shetty, R. (2005). "A novel implementation of transcription-based logic: Device design."

Smolke, C. D. and J. D. Keasling (2002). "Effect of gene location, mRNA secondary structures, and RNase sites on expression of two genes in an engineered operon." Biotechnol Bioeng 80(7): 762-76.

Weiss, R. (2001). "Cellular Computation and Communications using Engineered Genetic Regulatory Networks."

Whoriskey, S. K., M. A. Schofield, et al. (1991). "Isolation and characterization of Escherichia coli mutants with altered rates of deletion formation." Genetics 127(1): 21-30.

Yokobayashi, Y., R. Weiss, et al. (2002). "Directed evolution of a genetic circuit." Proc Natl Acad Sci U S A 99(26): 16587-91.

You, L., R. S. Cox, 3rd, et al. (2004). "Programmed population control by cell-cell communication and regulated killing." Nature 428(6985): 868-71.

7 Appendix

7.1 Microfluidic sortostat

I will design and operate a microfluidic system that enables more sophisticated screening and selection functions. Specifically, I will develop a microfluidic chemostat integrated with a cell sorter (i.e., a sort-o-stat). This microscope-based system will enable me to test whether or not more complicated screens and selections will be of practical use in service of evolving engineered biological systems.

Device Design

I have extended the design of a microfluidic chemostat developed by Frederick Ballagadde (Balagadde et al., 2005). The microfluidic chemostat has a 16nL reactor volume and can maintain 100 - 10^4 cells in semi-continuous, planktonic growth. Dilution in the microfluidic chemostat occurs in discrete events. Specifically, the reactor loop consists of 16 individually-addressable segments (Figure 9). These segments can be isolated, flushed with lysis buffer to remove old media and wall-growth, and then refilled with fresh media. The loop also contains a peristaltic pump in order to mix the cell population.

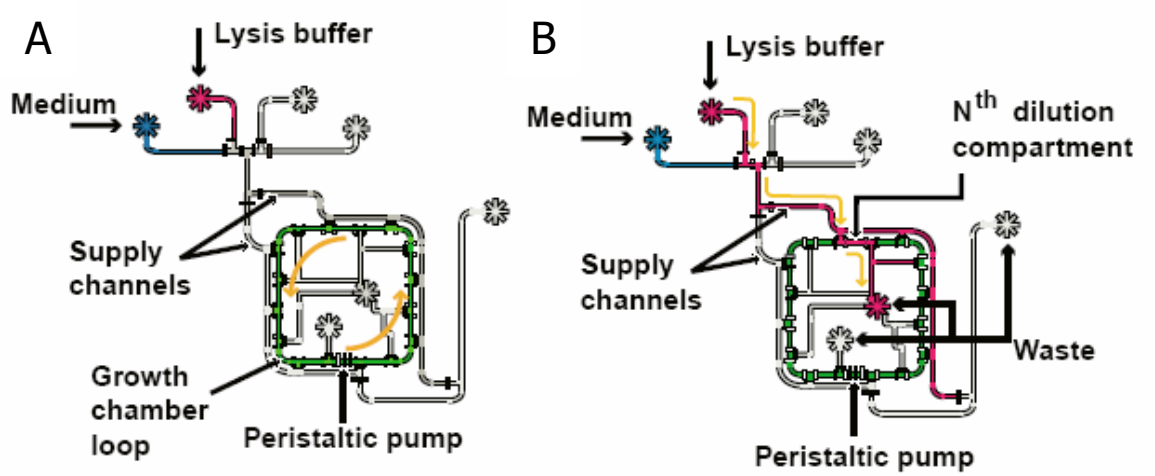

Figure 9: Schematic diagrams of a microfluidic chemostat. (A) Chemostat is operating in mixing mode. (B) Chemostat is operating in cleaning mode. Lysis buffer (red) is flowing from the lysis buffer inlet into one of 16 chambers in the chemostat and washing cells into the waste. Following this the chamber will be flushed with fresh media and the chemostat will return to mixing mode. (Balagadde et al., 2005)

I have extended the design by adding a 100x100um individually-addressable sorting chamber (Figure 10). This chamber is approximately 1/100th of the reactor volume and enables a subset of the population to be isolated, analyzed by quantitative microscopy, and then either flushed or allowed to remain in culture. In this way dilution events are made non-random and sophisticated selective pressures can be applied to the population on the basis of fluorescence output or cell morphologies.

Results

I designed the sorting extension for the original microfluidic chemostat design in AutoCAD (Figure 10) and ordered fabrication of the PDMS microfluidic chip from the Caltech Microfluidic Foundry [http://thebigone.stanford.edu/foundry/]. I have developed a microfluidic platform locally to operate the chip. The setup includes a Nikon TE2000 inverted microscope with an automated XYZ microscope stage, shutters, and filter cubes as well as automated manifolds for modulating pressure to actuate valves on the device. Additionally, the entire setup has been configured to be operated via LabView software and a custom application has been developed in LabView to operate the device, capture data, and sort cells automatically.

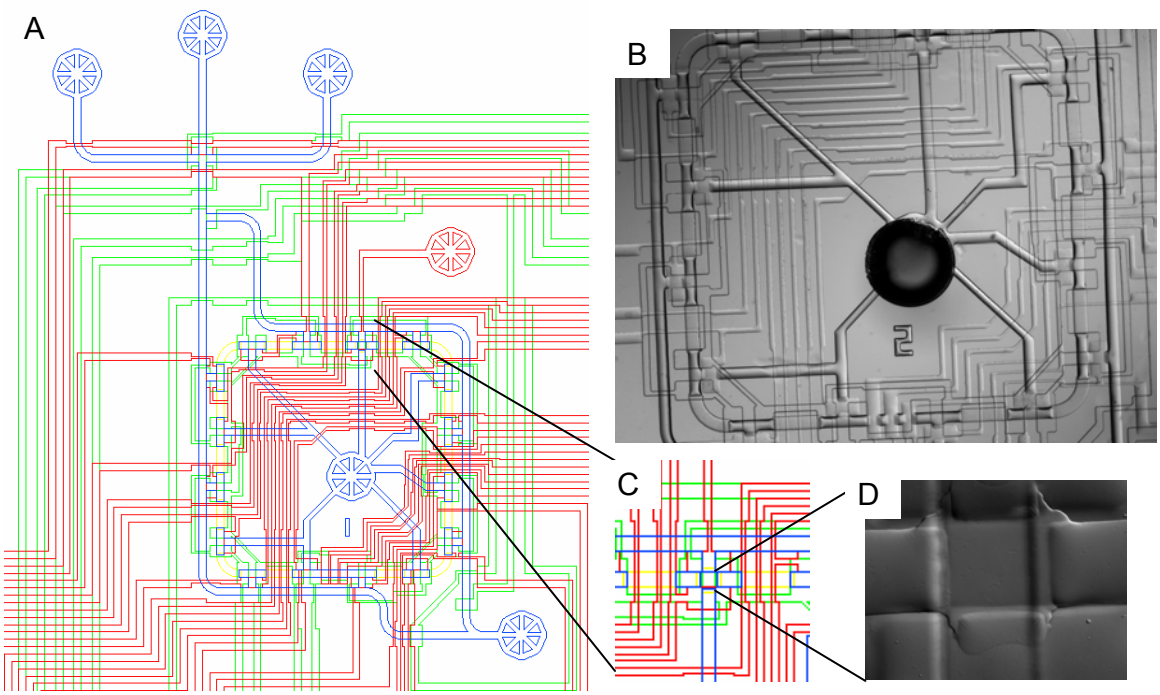

Figure 10: (A) AutoCAD design for a microfluidic sortostat. (C) and (D) outline the position of the sorting chamber. (B) Image captured on the microscope of the reaction loop and center waster channel.

I have successfully operated the device in normal chemostat mode (e.g., no sorting) for about 90 hours at dilution rates between 0.3-0.6 hr^{-1} (Figure 11). Next I will attempt to sort a mixed population of cells where one cell type expresses CFP and the other YFP. I hope to oscillate which population is selected against throughout the run and demonstrate that two strains can be maintained stably in the sortostat, something that cannot be done in a traditional chemostat.

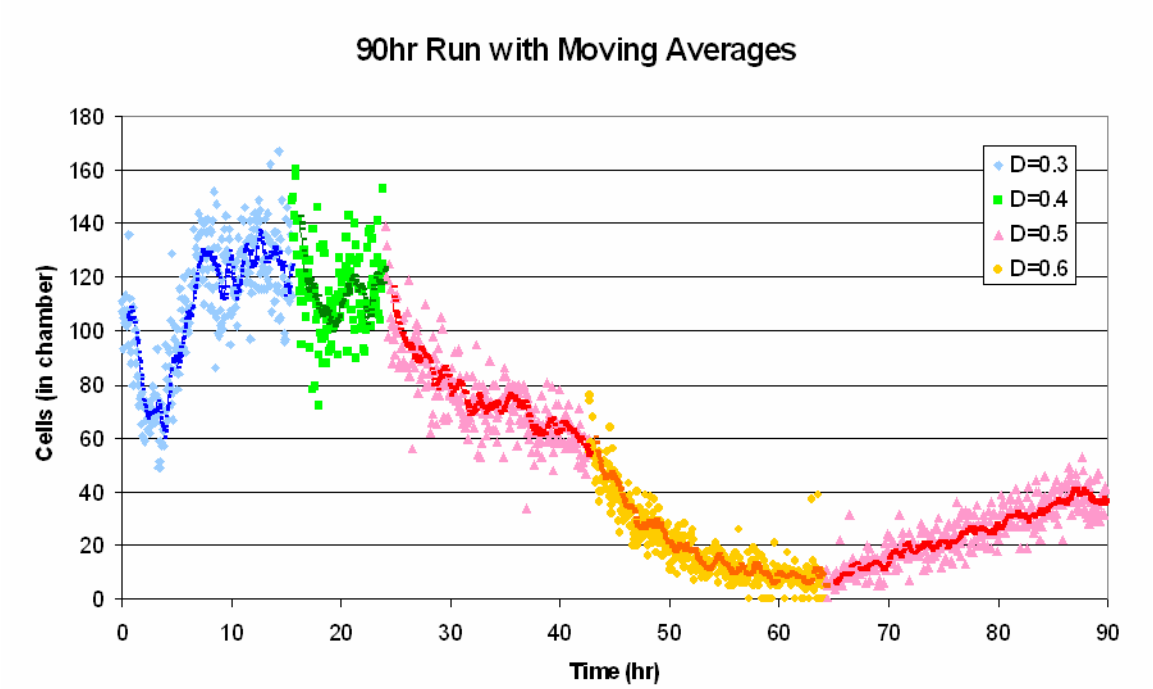

Figure 11: Cell counts in the sorting chamber for a 90hr sortostat run with device operating in normal chemostat mode, samples were captured at 3 minute intervals. The dilution rate was varied from $0.3hr^{-1}$ to $0.6hr^{-1}$ and culture density responded appropriately. The darker lines are a moving average with a ten time point window.